HENRI G. VETTER*
STEPHEN M. PIZER†
Massachusetts General Hospital
Boston, Mass. 02114

Perception of Quantum-Limited Images

The study suggests that the visual system approaches the problem presented by a quantum-limited image as a texture problem rather than a blurring of dots and an averaging of luminous intensities to treat it as an intensity-contrast problem.

ABSTRACT: *If a scene, illuminated by very few photons, is photographed, the result is a quantum-limited image, in which the arrival coordinates of the photons are individually recorded as bright dots and where no information is available between these dots. The nature of the analysis carried by the eye to reconstitute a continuous image when confronted with a quantum limited image, is explored and a model is developed, based on the assumption that the eye evaluates the distance between dots and their closest neighbors. This model permits a prediction of pictorial quality in good agreement with human preference.*

INTRODUCTION

DUE TO THE quantized nature of light, images obtained at extremely low illumination levels consist of a two-dimensional distribution of discrete events—the arrival points of the individual photons. If these events are recorded individually in the form of dots, we will call such a record a quantum-limited image (QLI), provided that the presentation of the image is such that the dots are individually resolved by the observer (see Figure 1 for example). Such images occur routinely in the fields of medical radioisotope scanning as well as in X-ray microprobe analysis.

In the field of radioisotope scanning, for example, the geometric distribution of a radioactive source is inferred from the recording of relatively few individual gamma rays, often no more than a few thousand for a whole picture. The primary record is thus a typical example of a QLI as defined above.

In such images the significant geometric features, such as boundaries between areas of different radioactivity levels, are deeply imbedded in statistical uncertainty due to the quantum-limited nature of the image. These boundaries must be inferred by the observer through some statistical analysis performed at the local level, involving sampling areas of significant extent. Thus the accuracy with which the original features can be recovered is limited not only by the resolution of the imaging device used, but also by the nature and efficiency of the statistical analysis performed by the visual system‡, together with the number of events available locally, which determines the statistical uncertainty level.

These two degrading factors are not independent because on any instrument it is possible to increase the number of events received at the expense of resolution. Thus a compromise must be made, and in order to be able to select the most favorable one, it is

* Biophysics Laboratory, New England Medical Center Hospitals, Boston, Mass.
† Department of Computer and Information Science, University of North Carolina, Chapel Hill, North Carolina.
‡ By *visual system* we mean all mechanisms in the eye and the brain involved in the imaging and interpretation of quantum limited images.

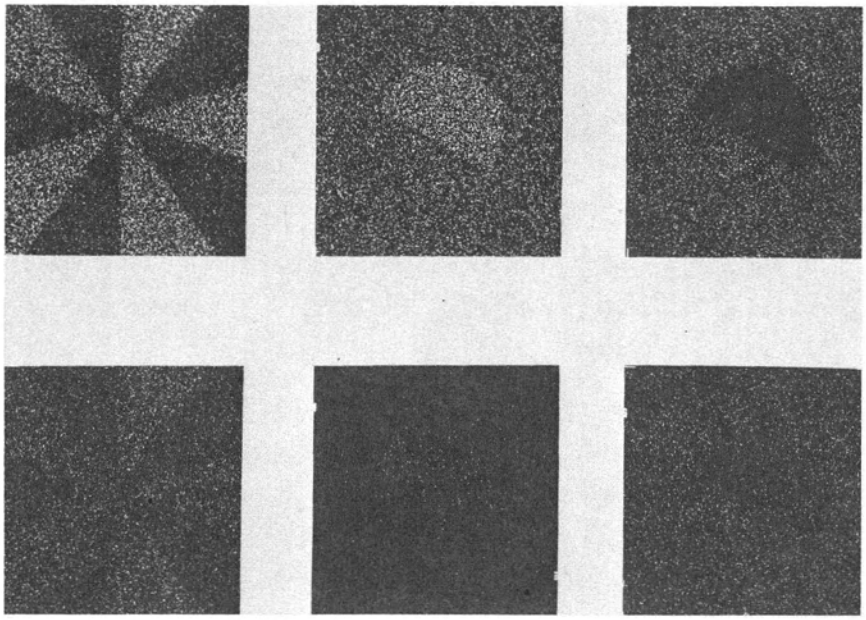

FIG. 1. Examples of Quantum-Limited Images used in this Study.

necessary to know how the visual system goes about analyzing random dot distributions to detect significant variations in the distribution parameters from one area to another. Thus this research deals with human perception of QLI's. Previous work done in this area[1,2] has focused on the determination of detection thresholds, and accordingly the stimuli used were very elementary. The results thus shed light on some individual aspects of the perception of QLI's but do not permit prediction of the over-all impression of quality conveyed by a specific QLI to an observer because the latter is due to the interplay, in an unknown fashion, between various factors such as ability for boundary extraction, *a priori* knowledge, etc.

Our specific concern here, on the other hand, is with this over-all impression of quality, and how it is affected jointly by the various parameters involved, above all dot population, contrast, and imaging resolution. In the absence of specific models for the joint effect of these parameters, not to mention the lack of an objective measure for quality, we are reduced to a black box, behavioral approach. The box will consist of a mathematical model of the process of human perception of QLI's which adequately approximates human behavior only in the sense that it will rate various versions of a picture in the same order in which human observers do, on the basis of subjective quality. The boot-strapping procedure we are compelled to follow often precludes absolute rigor in the arguments. We have tried to compensate for that by arguing our decisions from more than one point of view whenever possible. Only the ultimate performance of the model will justify our more arbitrary decisions.

CONTEXT FOR THE MODELS

The following is mainly intended to make plausible our assumptions about the process of perception of QLI's and should not be taken as claims to be validated. We consider a QLI as an intermediate record from which a continuous image is reconstituted by the visual system. Therefore, a measure of the pictorial quality of such an image must take into consideration not only the geometric (optical) degradation always introduced by an imaging device, but also the degradation due to the statistical analysis necessary to reconstitute the final continuous image from the set of individual quanta. The geometric degradation is solely controlled by the properties of the imaging device producing the quantum-limited image and is similar to that encountered with any optical imaging device. The statistical degradation, on the other hand, depends both on the sophistication of the statistical analysis *and* on the properties of the imaging device, which determine the number of dots over which the statistical analysis can be performed, and is peculiar to QLI's. The pro-

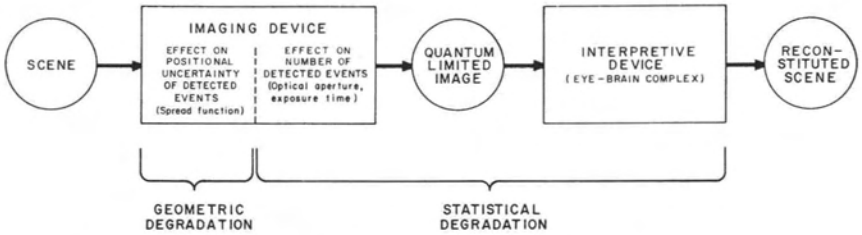

FIG. 2. Block diagram for perception of Quantum-Limited-Image.

cess we envisage is represented by Figure 2 which distinguishes clearly geometric and statistical degradations.

Accordingly, we can subdivide our objective in two. First and most, we will specify a model which will predict the *statistical quality* of a QLI and develop a suitable measure for it. Second, we will combine this measure with the imaging resolution and show that a measure of overall quality can easily be obtained, which is consistent with the preferences of a human observer. Unfortunately, lack of space will prevent us from going into the details of this latter step.

DEVELOPMENT OF STATISTICAL MODEL

In quantum-limited images the fine progression of gray levels which one can perceive in usual high intensity images has essentially disappeared, and thus we will assume that the information conveyed by a quantum-limited image is abstracted in the form of patterns formed by boundaries which are inferred by the visual system from significant changes in the results of statistical analysis performed over adjacent sample areas. That is to say, we assume that the predominant features are boundaries, perceived as discontinuities in the texture. These boundaries can be conceived as a sequence of independently perceived straight segments (elements), which must be long enough to permit the necessary statistical decision. The length of these boundary elements will determine the degree of detail which can be detected. Thus, the perceived quality will be determined by the length of the smallest boundary element detectable and the accuracy with which these elements can be located.

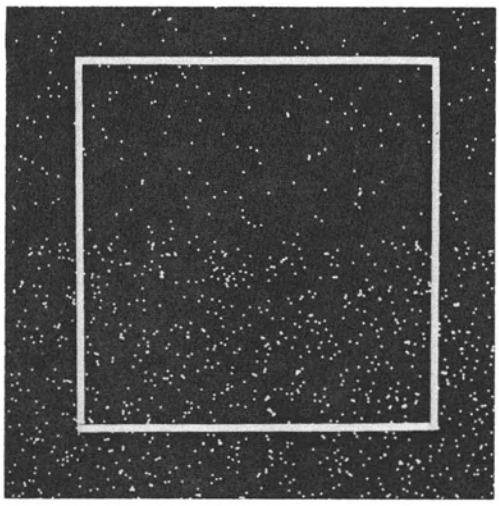

FIG. 3. Sampling element used for boundary perception.

Consider two adjacent random dot distributions with dot densities λ_1 and $\lambda_2 = \alpha\lambda_1$, where α is the contrast. As boundaries are not perceived directly but inferred through statistical analysis, the accuracy with which a boundary can be localized depends on the size of the sampling areas available in each of the two fields (Figure 3). The larger the area, the better the estimate of the statistical parameters and thus that of the location of the boundary. One must realize, however, that the use of a given sampling area precludes postulation of any other boundary element within that area. If we assume that the visual system aims at maximizing the number of perceivable boundary elements (without *a priori* knowledge of the pattern configuration), then the size of the sampling area will be automatically reduced to the point where the visual system just feels that a change is taking place within the area without being able to localize it further. That is to say that at this point one of the dimensions of the area will represent the length of the minimum boundary element perceivable while the other dimension will represent the uncertainty as to its transverse location.

The total sampling area will be essentially square because, *a priori*, there are no preferred directions in the pictures considered and thus the length of a boundary element (longitudinal resolution) should be of the same order as the uncertainty about its loca-

tion in the transverse direction. As both the accuracy with which the boundary can be defined and the length of the boundary element vary with the size of the sampling area, the pictorial quality will be monotonically related to a single index, the size of the sampling area. By analogy to optics we will attempt to define a distance, the *statistical resolution distance*, d_s, which will be proportional to the *effective diameter* of this area and thus be a suitable index to rate quantum limited images in terms of their informative value to the human eye.

Specific Models Considered

Three groups of models were considered, each embodying a different assumption as to the prevailing stimulus to the visual system. In group 1, the counting models, it is the number of dots present. In group 2, distance models, it is the average distance from a dot to its *closest* neighbor. In group 3, the area models, it is the *size of the holes* or empty areas between dots, for which the mean of the square of the distance between dots is a suitable index.

For each case the analysis involves reducing the sampling area until we reach a specified upper limit for the probability that an observed difference between the value of a particular index (e.g., number of dots) evaluated on either side of a presumed boundary might result by chance from a uniform random distribution. Making various assumptions specifying this composite uniform distribution leads to variations on each model identified by the subscripts 1, 2, and 3 in Table 1.

A summary of the model measures investigated is given in Table 1. We see that all the measures obtained are of the form $d_s = kf(\alpha)g(\lambda_1)$, where $g(\lambda_1)$ has in all cases the form $g(\lambda_1) = \lambda_1^{-1/2}$. Thus, we can discuss the models on the basis of the α dependence only (See Figure 4).

We note:

a. The curve corresponding to the counting model, is qualitatively different from all the other curves;

b. The area model is distinguished from the other two models by its sensitivity to the statistical assumptions, behavior which brings this model into question because we consider the essence of the model to be more in its physical basis than in its statistical assumptions.

Selection of Model

Procedure

How well do the models predict human evaluation of QLI's? To answer this question, sets of QLI's were produced by computer simulation.[3] They were printed in a 8×10-inch size, on high-contrast paper, white dots on black background. The dot size was kept constant and small compared to the average distance between dots. To permit a straightforward application of the model, each picture involved a single contrast. Thus, these quantum-limited images consisted of two-tone patterns where each tone was materialized by a random uniform dot distribution of a specified population density so that the distribution was discontinuous across a boundary. Such quantum-limited images are ideal ones in the sense that they could result only from a two-level scene with abrupt discontinuities imaged by an ideal device which does not introduce any geometric degradation.

In an attempt to eliminate the effect of a particular pattern on the results, the experiments were repeated with three different patterns. The first pattern, a ten-branch star (sectors), with branches alternately light and dark, is similar to typical test patterns used in optics (See Figure 1). In this pattern less populated and more populated fields play symmetric roles. The other two patterns consist of a half circle in a contrasting background. This pattern was selected as typical of those encountered in radioisotope activity mapping. It displays a range of boundary elements: straight, curved, and angular. Because of the non-symmetric nature of the

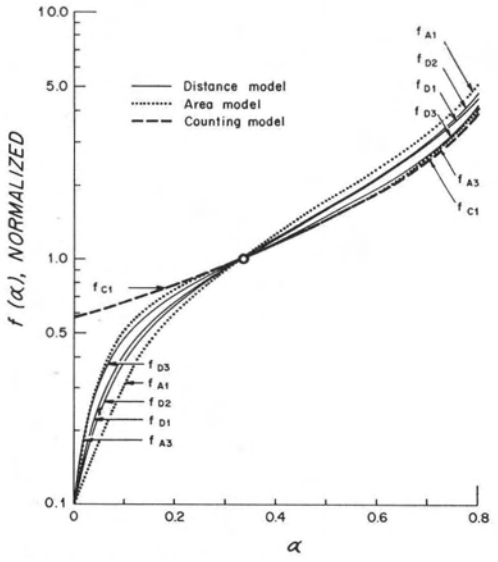

FIG. 4. Contrast dependence for various models (common point results from arbitrary normalization.)

TABLE 1. SUMMARY OF MODEL MEASURES

Statistical Assumptions		*Model Type*		
Subscript	*Description*	*Counting* (*subscript C*)	*Distance* (*subscript D*)	*Area* (*subscript A*)
1	σ of composite uniform distribution derived from uniform redistribution of dots	$d_{C1} = \dfrac{K(1+\alpha)^{1/2}}{\lambda_1^{1/2}(1-\alpha)}$	$d_{D1} = \dfrac{K\alpha^{1/2}(1+\alpha^{1/2})}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)}$	$d_{A1} = \dfrac{K\alpha}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)^{3/2}}$
2	σ of composite uniform distribution derived from index evaluated over whole sample area	$d_{C2} = d_{C1}$	$d_{D2} = \dfrac{K\alpha^{1/2}(1+\alpha^{1/2})^2}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)^{3/2}}$	$d_{A2} = d_{A1}$
3	σ of composite uniform distribution perceived directly from sample	$d_{C3} = d_{C1}$	$d_{D3} = \dfrac{K\alpha^{1/2}(1+\alpha^{1/2})\left(\dfrac{8}{\pi} - \dfrac{(1+\alpha^{1/2})^2}{1+\alpha}\right)^{1/2}}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)}$	$d_{A3} = \dfrac{K\alpha^{1/2}(1+\alpha^2)^{1/2}}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)^{3/2}}$

d = effective diameter of sampling area
λ_1 = dot density in denser region
α = contrast

pattern, it was realized in two sets, one with a more populated area on less populated background and the other with a less populated area on a more populated background.

Among the pictures of a set there are large qualitative differences: in some, with low contrast, the pattern is preserved thanks to a large number of dots; in others, with very few dots, the pattern is preserved thanks to high contrast. Thus the rating by an observer can only be qualitative; no *measure* of quality can be obtained from the subject but only a *preference* among the pictures expressed by a *ranking*.

The sets, consisting of about 24 pictures each, with contrasts varying between 10 to 1 and 3 to 2, and dot populations between 1,000 and 10,000 were submitted to a group of 15 human observers who ranked each set in order of pictorial quality.* These human rankings were then compared to the model rankings defined by the predicted index of pictorial quality, d_s, to determine the best matching model.

RESULTS AND INTERPRETATION

Having conflicting rankings by several subjects, we must first define the average human ranking or consensus, $\overline{H}$. Following Kendall, we will define it as the rank of the sum of the ranks[9]. It is obtained by adding for each picture the ranks it has scored in the various rankings and then rearranging the pictures on the basis of their total scores.

We must now compare the predictions of the different models with the consensus ranking $\overline{H}$. The model ranking is obtained by computing d_s for each picture in the set and ordering the pictures in increasing order of d_s. A measure of the agreement between models and human rankings is then the rank correlation coefficient τ as defined by Kendall.[9]

The τ values obtained by comparing the different model rankings to $\overline{H}$ for each set of pictures is given in Table 2 together with average values over the three sets of patterns. For comparison τ values measuring agreement among humans, and between individual human rankings H_i and the consensus $\overline{H}$ are also given.

Assuming the differences noted are significant, we conclude:

- The distance model is the best of those investigated† with all patterns and under all statistical assumptions.[1,2,3]
- Among the different versions of the distance model, d_{D1} and d_{D3} consistently give the best fit with little to choose between them except for questions pertaining to the internal consistency of the assumptions leading to d_{D3}**. Thus, we select d_{D1} as the statistical resolution distance.

$$d_s = Kd_{D1} = \frac{k\alpha^{1/2}(1 + \alpha^{1/2})}{\lambda^{1/2}(1 - \alpha)(1 + \alpha)} \qquad (1)$$

  where k is a constant scale factor which includes a human constant related to a confidence threshold level.
- The value d_{D1} fits the average human data as well as can be expected from any model as it fits as well as a human subject does on the average. In contrast, the counting models produce rankings which are difficult to reconcile with human rankings.

Two questions remain to be answered before the above conclusions can be taken seriously.

1. The significance of the differences observed in Table 2.
2. The sensitivity of our results to the particular measure of agreement which we have chosen.

Significance of Results Given in Table 2

Given that we accept the Kendall τ as a suitable measure of fit between model and human data, are the observed differences in τ significant, given the number of subjects used and the distribution of their answers?

We can answer this question with respect to any two models by asking how these differences would have been affected if more subjects had been included before calculating $\overline{H}$. Let us assume the most unfavorable and unlikely circumstances, in which all subsequent subjects would give rankings matching identically the model ranking dissenting more from the consensus. The effect of the inclu-

* The observers, all of whom were scientific personnel at Mass. General Hospital, were given the following instructions: "Consider the pictures as typical outputs from different cameras, and rank the pictures according to your preference if you had to select the best camera." By doing so, we hoped that, relying on the intelligence of the subjects, one could simplify the much more elaborate (if at all possible) experimental setup otherwise necessary to cancel out the effect of *a priori* knowledge of the pattern together with fortuitous arrangements of the dots in the various pictures. Otherwise we would have had to include both a multiplicity of pictures of "equivalent complexity" but presenting unpredictable patterns, and various statistical samples of each picture. Indeed, with a few extreme cases where we did repeat samples, the selection of the subjects was not importantly influenced by the differences between the samples.

† This result agrees with the results of Green et al.[1] indicating dot separation is an important factor for the human observer.

** Lack of space makes it impossible to discuss this matter. Details may be obtained from the authors.

TABLE 2. TABULATION OF HUMAN AND MODEL RANKINGS

Descriptive Contents	*Parameter*	*Numerical Values*			
		Pattern Used in Set			*Average over 3 Sets*
		Star (Sector)	*Dark Semicircle*	*Light Semicircle*	
—	Number of Rankings Contributed	15	16	13	—
—	Number of Pictures Ranked	24	25	25	—
Correlations within Human Rankings	$\bar{\tau}(\overline{H}; H_i)$ average	.918	.884	.917	.906
	$\bar{\tau}(H_i; H_j)$ average	.882	.839	.869	.863
Counting Model	$\tau(\overline{H}; d_{C1})$	.842	.833	.847	.841
Distance Model	$\tau(\overline{H}; d_{D1})$	.920	.887	.907	.905
	$\tau(\overline{H}; d_{D2})$	.890	.880	.893	.888
	$\tau(\overline{H}; d_{D3})$	.906	.900	.907	.904
Area Model	$\tau(\overline{H}; d_{A1})$	.825	.827	.840	.831
	$\tau(\overline{H}; d_{A3})$	.890	.887	.880	.886

sion of such subjects on the observed difference can be taken as a measure for the significance of this difference. For example, the number of d_{C1}-like subjects which must be added to reduce the difference between $\tau(\overline{H},d_{C1})$ and $\tau(\overline{H},d_{D1})$ to zero, compared to the number of real subjects (initial group), will be taken as a measure of the significance of the difference.

Figure 5 shows the upgrading and degrading of $\tau(\overline{H},d_{C1})$ and $\tau(\overline{H},d_{D1})$ respectively as a function of the number of d_{C1}-like subjects added. The crossover point occurs for about 8 added subjects, for an original population of 15. This result shows that, for our data, a high level of significance should be attached to the difference between 0.92 and 0.84 that is observed.

Using this result as a guideline to evaluate the other sets, we conclude that all the τ differences are significant except for those between distance models and d_{A1}. However, we have reservations with respect to the area model, because of its sensitivity to the statistical assumptions. Because its best version (d_{A1}) still performs more poorly than all distance models, we reject d_{A1} at this point.

SENSITIVITY OF RESULTS TO MEASURE OF FIT

Our data contains more information than just that given by rankings. In the case of the models, we have ignored the actual model measure and retained only the resulting rank. In the case of the human data, we have ignored the *strength of the consensus* indicated by the degree of unanimity observed between individual subjects. These arbitrary decisions could affect or restrict the meaningfulness of the results.

To deal more properly with the human data we can use the *normalized sum of the ranks* (NSR) defined as the sum of the ranks divided by the number of subjects. In the NSR the numerical distance between elements is, in a sense, a measure of the relative degree of

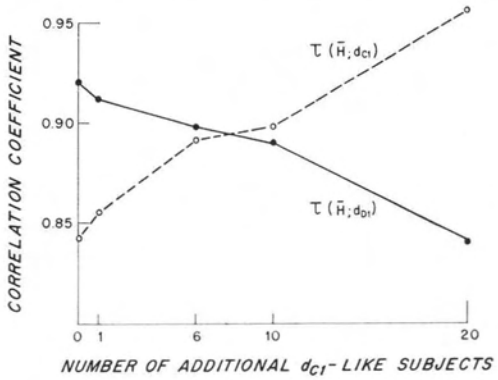

FIG. 5. Evaluation of significance of difference in rank-correlation coefficients.

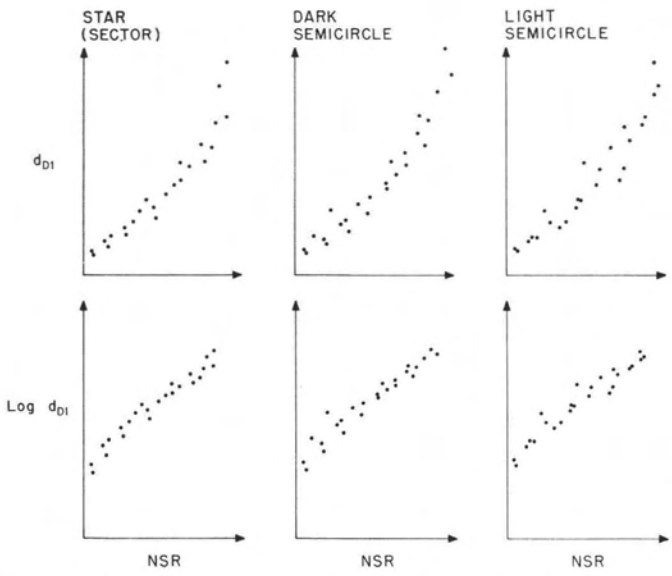

FIG. 6. Relationship between normalized sum of ranks (1) with d_s; (2) with log d_s.

unanimity between subjects in various portions of the ranking.

As far as the models are concerned, the way in which the quantitative information contained in d_s is used deserves comment. We want to use a monotonic function of d_s which gives the proper measure of *quality* of a QLI.

We have assumed quality Q is related to a distance, d_s (small distance—high quality). If we postulate that, as intuition suggests, equal *relative* changes in distance have equal perceptual significance, then ΔQ is proportional to $\Delta d_s/d_s$ so that Q is a constant, -log d_s, rather than a linear function of d_s.

Let us note here that though the argument above is not compelling, it is supported by the data. If one plots log d_s against ($-$NSR), which we have seen is our best objective estimate of pictorial quality, then a straight line fits the data very well (Figure 6).

This fit is shown by values obtained for the correlation coefficient between the NSR and log d_{DI}: .978, .969, and .962 respectively for the three patterns. These values are systematically higher than those obtained for the correlation coefficient between NSR and d_{DI} which are .947, .954 and .950. These results support the argument that if the measure of statistical resolution is d_s, the corresponding measure of subjective quality is a constant $-\log d_s$.

The quantitative information given by d_s (or constant $-\log d_s$) and the NSR can now be incorporated in various ways into our comparisons to produce various coefficients of fit, which represent a wide range of assumptions as to the nature of the information contained in the data. At one extreme is Kendall's τ which is based on the assumption that only qualitative distinction (rank) exists in the data. At the other extreme is the classical correlation coefficient ρ based on the assumption that the data contains truly quantitative information. Other coefficients represent attempts to take a hybrid point of view in the hope that it would be more realistic than either extreme.

A total of seven coefficients of fit* including Kendall's τ and the correlation coefficient have been calculated for all models and all sets of pictures. The results obtained are essentially independent of the coefficient used. This fact adds to the credibility of the results.

FINAL REMARKS AND COMBINED MODEL

The results of this study suggest that the visual system does approach the problem presented by a QLI as a texture problem rather than a blurring of dots and an averaging of luminous intensities to treat it as an intensity-contrast problem. Thus QLI's can be viewed as a particular variety of texture stimuli somewhat similar to those used by Julesz[4], but from which non-isotropic features are absent.

This suggests more work with isotropic textures, with distributions which, while appearing random to the unwarned observer,

* See note ** page 1184.

present internal regularizing constraints. Investigations with such distributions could help to single out which factors are extracted by the visual system to achieve the high degree of discrimination it displays.

For instance, we have observed that if an image is made up of fields of dots under a regularizing constraint, the observed improvement in picture quality over the random situation is at least qualitatively consistent with the predictions given by our distance model. We have also observed that a field of dots produced under a partially regularizing constraint appears systematically more populated to an unwarned observer, than a random field of equal population. Such an effect cannot easily be accounted for by a counting model and thus encourages us to believe in the distance model.

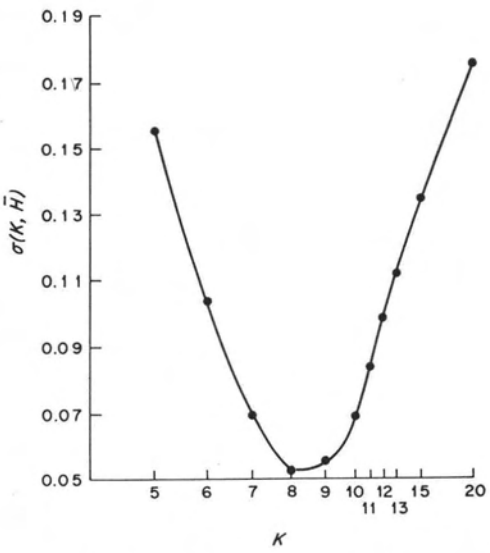

FIG. 7. $\sigma^2(k, \overline{H})$ as a function of k for all sets of pairs.

COMBINED MODEL

As said earlier, d_s, the measure resulting from this statistical model is to be combined with the measure of geometric degradation, d_g (see Figure 2) to produce an overall measure of degradation for a scan which is consistent with human evaluation of the scan. The geometric degradation encountered with any non-ideal imaging device is fully characterized by the device's point spread function. However, in the statistical context of QLI's the geometric degradation can be adequately characterized by a single index, the geometric resolution distance, d_g, defined as the diameter of the circle within which one half the flux of the two-dimensional spread function is contained. This measure has been shown to be reasonably independent of the shape of the spread function and to be in agreement with subjective human quality ratings[5].

The sampling area used by the eye for its statistical analysis also involves a degrading spread function, of unspecified shape, but characterized by d_s. Since spread function shapes are not crucial in this argument we will estimate the overall effect by the following formula

$$d^2 = d_s^2 + d_g^2 \qquad (2)$$

where d is the resulting resolution distance. This formula is rigorously valid in the case of gaussian functions.

Note that d_s has been determined only up to a scale factor, k, because of the relative nature of the ranking procedure used. A specific numerical value for k is needed, however, if Equation 2 is to be tested. The selection of k and the evaluation of Equation 2 were achieved as follows. A set of 23 pictures of the star pattern were generated by the simulation program with a triangular spread function with various d_g values. These geometrically *and* statistically degraded pictures (hybrid pictures) were submitted to subjects for rating. The subjects were given with each hybrid picture a set of progressively degraded statistical pictures of identical contrast (the same pictures used in the establishment of the statistical model.) They were asked to rate the hybrid pictures by inserting them at the most suitable place in the sequences of statistical pictures. A *human consensus* $\overline{H}$ was then determined defining empirical matches between statistical and hybrid pictures. For each pair, the corresponding quality measures were computed: $\log kd_{DI}$ for the statistical picture and $\log[(\mathrm{k}d_{DI})^2+(d_g)^2]^{1/2}$ for the hybrid picture, assuming various values of k.

The best value of k was selected as that minimizing the mean square difference between the two above measures, over all the pairs considered. Results are shown in Figure 7. The minimum occurs for $k=8$ and the sharpness of the minimum is an indication of the significance of the optimization of k. For $k=8$, model predictions do not differ more from the human consensus $\overline{H}$ than an individual does on the average. Figure 7 includes all contrasts tested. Detail anaysis, contrast by contrast, shows no systematic effect of contrast on the best value of k, and the variations observed (from 7 to 10) appear

quite consistent with the small size of the samples and the nature of the problem. Thus we consider this combined model to be adequate to predict the over-all quality of QLI's of the type considered. The complete formula representing this model is then

$$d^2 = d_0^2 + \left[\frac{8\alpha^{1/2}(1+\alpha^{1/2})}{\lambda_1^{1/2}(1+\alpha)(1-\alpha)}\right]^2. \qquad (3)$$

ACKNOWLEDGEMENTS

The authors are indebted to Drs. Gordon L. Brownell and Saul Aronow of the Physics Reaserch Laboratory at the Massachusetts General Hsopital for stimulating dissussions; to the scientific personnel of the laboratory for their cooperation in ranking endless sets of pictures; to Linda Boynton and Russell Carr of the same laboratory for their help with numerical computations; to Jacqueline Smith and Cathy Thompson for their generous seccretarial services.

This work was partially supported by U. S. Public Health Services Grant GM-889.

REFERENCES

1. B. F. Green and al. The detection of statistically defined patterns in a matrix of dots. *Amer. Jour. of Psychology 72:* 503 (1959).
2. A. Rose. The sensitivity performance of the human eye on an absolute scale. *J. Opt. Soc. Amer. 38:* 196 (1948).
3. S. M. Pizer. Simulation of Radiosotope Scans by computer. *Comm. ACM 9* (5): 358 (1966).
4. B. Julesz. Some recent studies in vision relevant to form perception, in *Models for the Perception of Speech and Visual Form.* W. Wathan-Dunn, ed. (MIT Press, Cambridge 1967), pp. 136–154.
5. H. G. Vetter. Characterization of geometric Imaging Properties of gamma-ray detectors. *Int. J. Appl. Rad. and Isotopes, 18:* 231 (1967).
6. H. G. Vetter and S. M. Pizer. A model for perception of quantum limited images and its application to radioisotope scanning, in *Proc. Seventh Int. Conf. on Medical and Biological Engineering.* Stockholm 5: 4 (1967).
7. G. L. Brownell. Theory of radioisotope scanning. *Proc. Second Int. Conf. on Med. Radioisotope Scanning,* IAEA, Athens, 1: 3 (1963).
8. S. M. Pizer and H. G. Vetter. Perception and Processing of Medical Radioisotope Scans, in *Pictoral Pattern Recognition.* Thompson Book Company, Washington, D. C. (1968). pp. 147–156.
9. M. G. Kendall, *Rank Correlation Methods,* Third Edition, Hafner Publishing Company, New York, 1962
10. M. G. Kendall and P. A. P. Moran. *Geometric Probability.* Charles Griffin & Co., Ltd., London 1962.